

VU Research Portal

Het experiment als paradigma voor interventie onderzoek in de fysiotherapie

Bouter, L.M.; Knipschild, R.

published in

Nederlands Tijdschrift voor Fysiotherapie
1987

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Bouter, L. M., & Knipschild, R. (1987). Het experiment als paradigma voor interventie onderzoek in de fysiotherapie. *Nederlands Tijdschrift voor Fysiotherapie*, 97(3), 56-61.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

HET EXPERIMENT ALS PARADIGMA VOOR INTERVENTIE-ONDERZOEK IN DE FYSIOTHERAPIE

L. M. Bouter en P. G. Knipschild

Terecht wordt van verschillende zijden op evaluatie van het effect van fysiotherapie bij verschillende aandoeningen aangedrongen. Dergelijk interventie-onderzoek dient daar waar mogelijk als een experiment te worden uitgevoerd. In het artikel wordt in hoofdlijnen het opzetten van therapeutische experimenten op het terrein van de fysiotherapie besproken. Een en ander wordt toegelicht met voorbeelden uit een recent onderzoeksvoorstel van de auteurs. Achtereenvolgens wordt stilgestaan bij het uitwerken van de probleemstelling, de componenten waaruit het effect van een behandeling bestaat en de noodzaak van randomiseren en blinderen. Tot slot wordt een tweetal indelingen van therapeutische experimenten besproken en wordt de lezer wegwijs gemaakt in meer specialistische literatuur over het opzetten van interventie-onderzoek.

In de kolommen van dit tijdschrift zijn in het recente verleden een aantal pleidooien gehouden voor een 'verwetenschappelijking' van de fysiotherapie (1, 2, 3). Hoewel verschillende auteurs binnen dit streven verschillende accenten leggen, is ook uit meer methodologisch getinte bijdragen duidelijk, dat evaluatie van het effect van fysiotherapie bij verschillende aandoeningen een hoge prioriteit heeft (4, 5, 6). Hetzelfde geldt voor de effecten van manuele therapie. Over de methodologie van dergelijk interventie-onderzoek gaat deze bijdrage.

Onder een interventie wordt in dit verband verstaan een gerichte poging om het natuurlijk beloop van een aandoening in gunstige zin te beïnvloeden. Een interventie kan dus zowel een voorziening, een voorschrift, een behandeling als een didactische mededeling zijn. Ook binnen de fysiotherapie worden allerlei verschillende interventies verricht. Interventie-onderzoek richt zich op het bestuderen van het effect van interventies. Dat betekent dat wordt getracht het instellen van de interventie te relateren aan veranderingen in één of meer effectparameters. In interventie-onderzoek is de centrale vraag of een therapie werkt, en niet hoe een eventuele werking tot stand komt. Inzicht in het werkingsmechanisme is derhalve noch het onderwerp van studie, noch een voorwaarde ervoor. Immers, voor het evalueren van het effect van een behandeling is het voldoende dat er consensus bestaat over het beoogde effect.

Dit artikel is als volgt opgebouwd. Allereerst wordt stilgestaan bij het formuleren van de probleemstelling. Vervolgens zal worden betoogd dat aan een experimentele aanpak de voorkeur dient te worden gegeven. Het experiment wordt daarbij opgevat als een paradigma. Deze door Kuhn (7)

drs. L. M. Bouter en Prof. dr. P. G. Knipschild, respectievelijk universitair docent en hoogleraar bij de vakgroep Epidemiologie van de Rijksuniversiteit Limburg.
Correspondentieadres: Vakgroep Epidemiologie, Rijksuniversiteit Limburg, Postbus 616, 6200 MD Maastricht.

in de moderne wetenschapsleer geïntroduceerde term wordt hier gebruikt in de betekenis van ideaalvoorbeeld. De methodologie van het therapeutisch experiment staat centraal in de rest van het betoog. De nadruk zal daarbij liggen op de verschillende theoretische en praktische keuzes waarvoor de onderzoeker zich geplaatst ziet. Eén en ander zal worden geïllustreerd met een onlangs geformuleerd voorstel voor een interventie-onderzoek naar het effect van fysiotherapie en manuele therapie bij niet-specifieke rug- en nekklachten. In dit onderzoeksvoorstel wordt beschreven hoe op experimentele wijze het effect van verschillende behandelingen bij niet-specifieke rug- en nekklachten kan worden geëvalueerd. Behalve fysiotherapie en manuele therapie, zal ook behandeling door de huisarts in het onderzoek worden betrokken. De redenen hiervoor zullen hieronder worden toegelicht. Het onderzoeksvoorstel is als illustratie gekozen in de hier gepresenteerde uiteenzetting, omdat het één van weinige concrete voorbeelden van interventie-onderzoek binnen de Nederlandse fysiotherapie is. Uiteraard kan in dit bestek aan alle nuances van het onderzoeksvoorstel geen recht worden gedaan. Naar verwachting zal het voor de lezer niet moeilijk zijn om zelf een groot aantal andere voorbeelden van interventie-onderzoek in de fysiotherapie te bedenken.

Probleemstelling

Een heldere beschrijving van het probleem waarop het onderzoek beoogt een antwoord te geven is van eminent belang. In geval van interventie-onderzoek zal de centrale vraag veelal luiden of een bepaalde nieuwe interventie effectiever is dan de bestaande behandeling, een placebo of geen behandeling. In het gekozen voorbeeld gaat het om het probleem dat langdurige rug- en nekklachten veelvuldig in de huisartspraktijk worden gepresenteerd en dat de etiologie en de pathogenese van deze klachten veelal niet bekend is. Er zijn voor deze grote groep patiënten verschillende behandelingen mogelijk, waarvan nog onvoldoende duidelijk is of, en in welke mate deze werkzaam zijn.

Bij het uitwerken van de probleemstelling van een interventie-onderzoek dient een drietal vragen te worden beantwoord.

1. Welke interventies zullen met elkaar worden vergeleken?

Veelal wordt een recent geïntroduceerde interventie (geneesmiddel of een ander soortige behandeling) vergeleken met de gebruikelijke behandeling. Zo wordt in het voorbeeld manuele therapie vergeleken met fysiotherapie en met behandeling door de huisarts. Omdat van alle drie deze behandelingen de specifieke werkzaamheid nog onvoldoende vaststaat, is het bovendien van belang om na te gaan of deze effectiever zijn dan een placebo-behandeling. Uiteraard is het steeds zaak om vóór het onderzoek de te vergelijken interventies zo goed mogelijk af te bakenen, en om tijdens het onderzoek precies te beschrijven waaruit de daadwerkelijk uitgevoerde interventies bestonden. De lezer van de resulterende publicatie weet dan wat er is bedoeld en kan deze interventies desge-

wenst in de eigen patiëntenzorg of in een ander onderzoek toepassen.

2. Bij wie zullen de interventies worden uitgevoerd?

Het gaat hier om de samenstelling van de onderzoekspopulatie. Wie komen in het kader van de probleemstelling in aanmerking om in het onderzoek te worden toegelaten, is hier de vraag. Behalve criteria voor toelaatbaarheid, dient ook te worden aangegeven, uit welke bron de patiënten zullen worden betrokken. In het voorbeeld wordt dit als volgt omschreven. Het gaat om patiënten (N=400) die minstens twee maanden niet-specifieke rug- of nekklachten (pijn en/of bewegingsbeperking) hebben, die reproduceerbaar blijken te zijn in actief en passief bewegingsonderzoek. Met niet-specifiek wordt bedoeld dat de huisarts op grond van zijn of haar normale beleid geen evidente oorzaak (maligniteit, osteoporose en dergelijke) heeft aangetoond. Dergelijke patiënten komen voor het onderzoek in aanmerking, voorzover ze zich in een bepaalde periode vervoegen bij één van de twintig in het onderzoek participerende huisartsen in Zuid-Limburg. In deze omschrijving valt op dat het bestudeerde spectrum van de aandoeningen op een aantal wijzen is ingeperkt. Vaak worden ook restricties aan leeftijd, geslacht of andere prognostisch relevant geachte variabelen gesteld. De onderzoekspopulatie dient homogeen te zijn. Dit streven wordt begrensd door de klinische relevantie van de gehanteerde criteria voor een bepaald indicatiegebied, en de mogelijkheid om nog voldoende van dergelijke patiënten te kunnen vinden.

3. Welke zijn de relevante effectmaten?

De kwestie is hier in welke maat het effect zal worden uitgedrukt. Het is zaak daarbij te kiezen voor de maten die het meest relevant zijn voor het gezondheidsprobleem dat aan de orde is (10). De belangrijkste uitkomst die de interventie beoogt staat daarin centraal. En niet in de eerste plaats de parameter die het meest objectief en met grote precisie kan worden vastgesteld. De volgorde is steeds dat eerst in conceptuele termen wordt aangeduid wat de belangrijkste effectmaten zijn. Pas daarna wordt bij de uitwerking van het onderzoeksvoorstel naar operationalisaties van deze maten gezocht. Een ander

punt is de bepaling van het moment of de momenten waarop de effectmeting dient plaats te vinden. Ook hier is een bezinning op de ideale situatie geboden. Vervolgens kan worden getracht om deze in de onderzoeksopzet zo goed mogelijk te benaderen. In het voorbeeld zijn pijn en onderhouden lichamelijke beperkingen in het dagelijks functioneren als belangrijkste effectmaten gekozen. Daarnaast zal ook de beweeglijkheid van de wervelkolom als effectparameter dienen. Als operationalisatie van deze effectmaten wordt zoveel mogelijk gekozen voor internationaal erkende en gestandaardiseerde meetinstrumenten. Voor zover die er niet zijn, worden de meetinstrumenten vóór het onderzoek ontwikkeld en getest. De patiënten in het onderzoek worden op verschillende momenten vanaf het begin van de behandeling, tot maximaal een jaar na aanvang, aan een effectmeting onderworpen.

Waargenomen effect

Nu de probleemstelling is geformuleerd en uitgewerkt, is de volgende stap het kiezen van de onderzoeksopzet. Daarbij is het goed te bedenken dat het waargenomen effect (WE) van een interventie in principe is opgebouwd uit vier verschillende onderdelen (zie schema 1). De achtergrond hiervan is dat de interventie slechts één van de determinanten is die de effectparameters beïnvloeden.

1. Het specifiek effect (SE) van de interventie

Doorgaans is de onderzoeker vooral hierin geïnteresseerd. Helaas is deze factor niet direct meetbaar. Gelijktelling met het waargenomen effect is slechts in zeer uitzonderlijke omstandigheden geoorloofd. Toch is een dergelijke gelijkstelling de impliciete achtergrond van menige gepubliceerde patiëntenserie. In ons voorbeeld zou een op zichzelf staande constatering dat de pijn bij met manuele therapie behandelde patiënten in zes weken met 40% afnam, een weinig informatieve mededeling zijn. Er staat daarmee immers allerm minst vast dat de pijn ten gevolge van de manuele therapie afnam.

2. Het natuurlijk beloop (NB) van de aandoening

Hoewel niet alle patiënten en (para)medici zich dit altijd ten volle lijken te realiseren, is er ook verbetering mogelijk zonder dat van enige interventie sprake is. Vanzelfsprekend behoort ook een verslechtering bij afwezigheid van interventie tot de mogelijkheden, al dan niet met fatale afloop of een resterende chronische aandoening. Omdat maar weinig aandoeningen tegelijkertijd wel bestudeerd maar niet behandeld worden, is over het natuurlijk beloop van een aandoening door de tijd veelal weinig bekend. Het zal duidelijk zijn dat verschillen in natuurlijk beloop tussen de patiënten die aan de in het onderzoek vergeleken interventies worden onderwor-

pen, de waargenomen effecten in belangrijke mate zullen beïnvloeden. Het is daarom van groot belang dat de prognose van de patiënten bij alle interventies hetzelfde is. In het voorbeeld zou het daarom verkeerd zijn om meer patiënten met een gunstige prognose aan de manueel therapeut toe te wijzen.

3. De externe factoren (EF) die het effect van de interventie beïnvloeden

Een waargenomen effect kan, zoals gezegd, in principe behalve door de interventie, ook door één of meer andere determinanten worden beïnvloed. Deze andere determinanten zijn versturende variabelen (confounders), in de zin dat ze het zicht op de relatie interventie-effect vertroebelen. Voor zover deze (nog) niet bekend zijn, vallen ze onder het natuurlijke beloop. De invloed van confounders die al wel als zodanig bekend zijn, kan in zekere mate door een elegante onderzoeksopzet worden geëlimineerd. Daarom heeft het zin dergelijke confounders afzonderlijk aandacht te geven. Dit geschiedt veelal onder de noemer 'externe factoren'.

Veelvuldig terugkerende voorbeelden hiervan zijn: leeftijd, geslacht, therapietrouw (compliance), periode waarover de klachten bestaan, voorafgaande behandeling en medicatie of therapie naast de onderzochte interventie. Voor een aantal aandoeningen is ook een invloed van bepaalde leefgewoonten op het beloop van de aandoening en daarmee op de effectmaten bekend. In het voorbeeld worden ondermeer de confounders leeftijd, geslacht, localisatie van de klachten, gebruik van pijnstillers en compliance in de beschouwing betrokken.

Tot de externe factoren behoort eveneens het zogenoemde placebo-effect van de behandeling. Hiermee wordt het fenomeen bedoeld dat iedere interventie behalve een eventueel specifiek effect, ook een specifiek effect heeft. Dit effect zal zich in pure vorm manifesteren bij patiënten die een placebo-geneesmiddel of 'sham-treatment' krijgen. Het gaat dan om een voor de patiënt geloofwaardige interventie waarvan, zonder dat de patiënt dat weet, het veronderstelde werkzame bestanddeel is weggelaten. Hoewel over het werkingsmechanisme van het placebo-effect nog maar weinig bekend is, is het goed om te benadrukken dat de invloed ervan, ook op de meest 'harde' effectparameters, even 'echt' is als die van de specifieke componenten van de interventie. In het voorbeeld wordt een kwart van de patiënten onderworpen aan een placebo-fysiotherapie (ondermeer ultrageluid zonder energieafgifte) en placebo-manuele therapie (uitvoerige fysische diagnostiek en enkele onschuldige handgrepen).

4. De meetfouten (MF) bij de effectmeting
Deze component berust op de overweging dat het waargenomen effect niet uitsluitend

Schema 1. (Toelichting in de tekst)

$$WE_i = SE_i + NB_i + EF_i + MF_i$$

$$WE_r = SE_r + NB_r + EF_r + MF_r$$

$$WE_i - WE_r = SE_i - SE_r$$

$$\begin{aligned} \text{alleen indien: } NB_i &= NB_r \\ EF_i &= EF_r \\ MF_i &= MF_r \end{aligned}$$

tend op de vorige drie componenten (SE, NB en EF) berust. Er kunnen immers ook bij het meten van de effectparameters fouten binnensluipen. Met name wanneer de optredende fouten samenhangen met de interventie van de patiënt, kan de resulterende verstoring aanzienlijk zijn. Dit is vooral het geval wanneer de effectmeting een subjectief oordeel van de patiënt of een observator behelst en de preferentie voor (of afkeer van) een bepaalde interventie dit oordeel beïnvloeden kan. De oplossing van dit probleem is ervoor zorgen dat de gene die de effectmeting verricht, niet weet welke interventie werd gedaan. Helaas is dit lang niet altijd te realiseren. In het voorbeeld lukt dat ten aanzien van de patiënt hopelijk voor de placebo-therapie, maar in ieder geval niet voor de behandeling door de eigen huisarts. De observator die de beweeglijkheid van de wervelkolom evalueert, zal niet weten welke interventie bij de betreffende patiënt werd uitgevoerd.

Therapeutisch experiment

Zoals gezegd, richt interventie-onderzoek zich op het specifieke effect van een behandeling. Aan de orde is interventie als specifieke oorzaak van afname van de klachten. Daartoe dienen de andere drie groepen determinanten (NB, EF en MF) onder controle te worden gehouden. In schema 1 is aangegeven dat het verschil in specifiek effect tussen de index (i) en de referentie (r) behandeling gelijk is aan het verschil in waargenomen effect. Dit op voorwaarde dat het natuurlijk beloop, de invloed van externe factoren en de meetfouten bij de index en de referentie groep hetzelfde zijn. Deze voorwaarde moet worden gerealiseerd door keuze van de onderzoeksopzet. Waar dat uitvoerbaar is, biedt het therapeutisch experiment (Randomized Controlled Trial) verreweg de beste mogelijkheden om de gewenste controle over de determinanten van de effectparameter te bewerkstelligen. Interventie onderzoek dat niet als een therapeutisch experiment is opgezet, wordt tegenwoordig in brede kring als zeer discutabel beschouwd. Dat komt doordat het ontbreken van controle over de determinanten aanleiding kan zijn tot een ernstige vertekening van de onderzoeksresultaten. In het therapeutisch experiment wordt deze controle uitgeoefend met behulp van randomisering en blinding.

Randomisering

Cruciaal voor de geldigheid (interne validiteit) van interventie-onderzoek is de wijze waarop de vergeleken interventies aan de patiënten in het onderzoek worden toegewezen. Medici en para-medici hebben doorgaans het beste met de patiënt voor: men zal bij voorkeur aan een patiënt die behandeling toeschrijven, waarbij men de kans op herstel het grootste acht. Een arts of fysiotherapeut zal derhalve in het kader van een interventie-onderzoek geneigd

zijn om de interventie van zijn of haar voorkeur aan patiënten met een relatief gunstige prognose toe te bedelen. Hierdoor treedt een ernstige vertekening op van de onderzoeksresultaten: 'confounding by indication' (11, 12). Immers, verschillen in effect hangen nu niet meer alleen met de toegepaste interventie samen, maar ook met verschillen in prognose tussen patiënten die aan een bepaalde interventie werden toegewezen. Van 'confounding by indication' kan sprake zijn in alle gevallen waarin het mogelijk is dat de ingeschatte prognose de kans maakt om aan een bepaalde interventie te worden toegewezen beïnvloedt. In het voorbeeld zou een dergelijke vertekening optreden wanneer huisartsen vooral die patiënten naar een manueel therapeut zouden verwijzen, waarbij de rug- of nekklachten al langer bestaan en waarbij eerdere therapie zonder succes bleef. Het zal duidelijk zijn dat hier geen boze opzet in het spel hoeft te zijn; ook onbewust kan een ernstige vorm van 'confounding by indication' tot stand komen.

De standaardoplossing van dit probleem is het toewijzen van de verschillende interventies door het lot. De voorkeuren van patiënten en (para-)medici zijn daarmee buiten spel gezet. Dit is de essentie van randomisering. Daarbij is het zaak zorgvuldig te bewaken dat de toewijzing ook werkelijk 'at random' plaatsvindt en dat op het moment van randomisering de toegewezen interventie onveranderbaar wordt vastgelegd. Behalve het gelijkelijk distribueren en daarmee controleren van de determinanten besloten in het natuurlijk beloop, bewerkstelligt de randomisering hetzelfde voor de externe factoren (EF) en voor de meetfouten (MF) voor zover deze niet systematisch zijn. Het is om deze reden dat randomisering als een panacee voor potentiële vertekening wordt gepropageerd en kan worden opgevat als het meest essentiële bestanddeel van een therapeutisch experiment. Het is echter een hardnekkig misverstand te menen dat randomisering garandeert dat de diverse determinanten ook daadwerkelijk gelijkelijk over de verschillende interventies zullen verdeeld. Immers, ongelijke verdelingen kunnen door het toeval wel degelijk ontstaan. Met name in onderzoeken met weinig patiënten kan een dergelijke ongelijke verdeling gemakkelijk optreden en grote problemen geven bij de analyse van de onderzoeksresultaten.

Ten aanzien van de determinanten (EF) waarvan het bestaan bekend is, kan men achteraf in de analyse nagaan in hoeverre deze gelijkelijk over de interventie groepen waren verdeeld. Een ongelijke verdeling van dergelijke determinanten behoeft geen ramp te zijn en vormt niet noodzakelijkerwijs een bedreiging voor de geldigheid van de uitkomsten. Er zal dan een zogeheten gestratificeerde analyse worden uitgevoerd, waarbij steeds voor één categorie van de ongelijke verdeelde determi-

nant naar het effect van de verschillende interventies wordt gekeken. Vervolgens wordt van dit effect een gewogen gemiddelde over de op deze wijze ontstane strata berekend. In ons voorbeeld zouden door het toeval bij de randomisering veel meer mannen aan de fysiotherapie en veel meer vrouwen aan de manuele therapie kunnen worden toebedeeld. Wanneer nu tevens kan worden aangenomen dat bij vrouwen rug- en nekklachten veel sneller genezen, zal er sprake zijn van vertekening. Omdat de betreffende determinant bekend is en voor alle patiënten in het onderzoek is vastgesteld, is in dit geval correctie met behulp van een gestratificeerde analyse mogelijk. Dit houdt in dat het effect van de bestudeerde interventies eerst voor vrouwen en mannen afzonderlijk zal worden berekend. Vervolgens wordt dan over deze twee strata een (gewogen) gemiddeld effect berekend.

Uit statistisch oogpunt is een dergelijke ongelijke verdeling van externe factoren (confounders) weinig efficiënt. Daarom probeert men veelal het lot een handje te helpen. Stratificatie vooraf garandeert dat de confounders waarvoor deze procedure wordt ondernomen gelijkelijk over de interventie groepen zullen zijn verdeeld. In het voorbeeld zou er dan onder de vrouwen en onder de mannen afzonderlijk worden gerandomiseerd. Een andere toepassing van dit principe is het afzonderlijk uitvoeren van de randomisering in de vier subregio's van het onderzoeksgebied. De reistijd van patiënten blijft daardoor beperkt, terwijl toch manuele therapie, fysiotherapie, placebo-therapie en behandeling door de huisarts in iedere subregio in gelijke mate aan bod komen. Het zal duidelijk zijn dat een dergelijke stratificatie vooraf slechts voor een beperkt aantal externe factoren mogelijk is. Het aantal verschillende strata wordt daarbij bepaald door het aantal externe factoren dat bij de stratificatie vooraf betrokken wordt, en door het aantal categorieën per factor. In het voorbeeld zijn er nu $2 \times 4 = 8$ strata gevormd van $400/8 = 50$ patiënten. In theorie is het mogelijk om door te gaan met stratificatie vooraf, totdat er in ieder stratum voor elke behandeling nog maar één individu beschikbaar is. Dit wordt individuele matching genoemd.

Voor de confounders die onbekend of wel bekend maar niet gemeten zijn, is stratificatie vooraf niet mogelijk. Voor de gelijke verdeling van deze factoren over de interventie groepen moet de randomisering zorgen. Voor deze confounders is ook stratificatie (achteraf) in de analyse niet mogelijk, zodat vertekening ingeval van toevallige ongelijke verdeling in principe mogelijk blijft. In feite vormen deze onbekende confounders één van de belangrijkste redenen tot randomiseren. Een ander argument tot randomiseren is gelegen in het feit dat belangrijke confounders soms niet of moeilijk nauwkeurig meetbaar zijn. Noodzake-

lijkerwijs levert in dat geval zowel stratificatie vooraf als stratificatie achteraf een gebrekkige controle op. In het voorbeeld vormen psychologische factoren zo'n moeilijk meetbare confounder. Er zijn redenen om aan te nemen dat dergelijke factoren zowel in het ontstaan als in het beloop van rug- en nekklachten een rol spelen. Omdat de relevante psychologische kenmerken zich slechts gebrekkig laten meten, moet voor een gelijke verdeling over de onderzoeksgroepen op randomisering worden vertrouwd.

Het interventie-onderzoek begint in feite op het moment van randomisering. Eénmaal gerandomiseerd, definitief in de onderzoekspopulatie is het principe. Ook is het van cruciaal belang dat de patiënt, wat er ook gebeurt, in de analyse wordt meegenomen als behorende bij de interventie die bij de randomisering was toegewezen. Afwijking van dit 'intention to treat' principe kan aanleiding geven tot serieuze vertekening van de onderzoeksresultaten en doet in feite de voordelen van randomisering weer deels teniet. Zo zou in het voorbeeld de situatie kunnen ontstaan dat de therapietrouw bij de behandeling door de fysiotherapeut geringer is dan bij de andere behandelingen. Wanneer dit komt omdat patiënten terecht vinden dat fysiotherapie de klachten niet verhelpt en deze afvallers niet in de analyse worden betrokken, ontstaat een vertekening van de onderzoeksresultaten waarbij het effect van fysiotherapie te hoog wordt ingeschat.

Blindering

Hoewel randomisering in belangrijke mate controle over de determinanten tot stand brengt, zijn er vormen van vertekening die hiermee niet zijn te voorkómen. Voor een causaal interpreteren van een verschil in waargenomen effect tussen de interventies is het een voorwaarde dat zowel het effect van de externe factoren als van de meetfouten bij de verschillende interventies vergelijkbaar zijn. Aan deze voorwaarde is niet voldaan wanneer de aard van de interventie van invloed is op het effect van de externe factoren of op de meetfouten. Het gaat hier met name om verschillen in placebo-effect en verschillen in de perceptie van het effect tussen verschillende interventies. In het voorbeeld zal wellicht de wetenschap dat men een 'nep-behandeling' krijgt, het placebo-effect van deze interventie tot nul reduceren. Het gevolg is een onderschatting van het placebo-effect van behandeling door fysiotherapeut, manueel therapeut en huisarts. Een ander probleem ontstaat wanneer de observator van de beweeglijkheid van de wervelkolom een groot voorstander is van manuele therapie en weinig ziet in de andere interventies. Het gevolg daarvan is dat die observator het effect van manuele therapie wellicht rooskleuriger zal percipiëren. Hierdoor ontstaat wederom vertekening van de onderzoeksresultaten.

Het antwoord op deze potentiële vormen van vertekening is blindering. Daarbij gaat het erom te maskeren welke patiënt welke interventie krijgt. Blindering is mogelijk voor de patiënt, voor de behandelend (para-)medicus, voor de observator van het effect en voor de analyser van de data. Het ideaal is volledige blindering op deze vier nivo's. Alleen degene die na de data-analyse een verslag of onderzoeksartikel schrijft, weet dan na afloop van de trial wie waarmee werd behandeld. Dit ideaal is alleen in geval van geneesmiddelenonderzoek volledig realiseerbaar. Daarbij speelt de placebo een cruciale rol. Dat is een 'geneesmiddel' dat er voor alle betrokkenen in ieder opzicht hetzelfde uitziet, maar waarin de ingrediënten waaraan de specifieke werking wordt toegeschreven ontbreken. Het placebo-effect van de vergeleken interventies (inclusief de placebo-interventie zelf) zal dan gelijk zijn en geen van de betrokkenen zal weten wie welke interventie kreeg.

Daar waar de interventie niet uit een geneesmiddel bestaat, is volledige blindering op alle vier de nivo's vrijwel nooit haalbaar. Dan is de richtlijn dat men tracht blindering door te voeren voor zover mogelijk. Daarbij dient zo goed mogelijk te worden nagegaan in hoeverre de blindering daadwerkelijk tot stand kwam. Wanneer de blindering onmogelijk is, of blijkt te zijn mislukt, zal moeten worden nagegaan in hoeverre hierdoor de onderzoeksresultaten kunnen zijn vertekend. Bij de presentatie en interpretatie van het onderzoek dient hiermee vervolgens rekening te worden gehouden. In het voorbeeld is blindering van de therapeut niet mogelijk. Men weet immers of men fysiotherapie, manuele therapie of placebo-therapie bedrijft. Het is daarom zaak ervoor te zorgen dat alle interventies met een vergelijkbaar enthousiasme en overtuigingskracht worden uitgevoerd. Blindering van de patiënt is in het voorbeeld slechts gedeeltelijk mogelijk. Ten aanzien van de placebo-therapie is de blindering cruciaal. Het bestaan van de placebo-therapie zorgt op zijn beurt weer voor (partiële) blindering van de fysiotherapie en de manuele therapie. In het voorbeeld is geen sprake van blindering van de patiënt voor de door de eigen huisarts verstrekte therapie. De observator van de effectmaten is in het voorbeeld ten dele de patiënt zelf en ten dele de projectmedewerker die de beweeglijkheid van de wervelkolom vaststelt. Laatstgenoemde persoon kan weer wel volledig worden geblindeerd en dit is ook wenselijk. Patiënten zullen steeds vóór het contact met deze observator worden geïnstrueerd om niets over de ontvangen behandeling te vertellen. Vooral effectparameters waarbij van de observator een subjectieve beoordeling wordt gevraagd, zullen gevoelig zijn voor vertekening ten gevolge van het ontbreken van blindering. Degene die de onderzoeksresultaten analyseert kan volledig worden geblindeerd door de toegepaste behande-

lingen zo te coderen dat ze niet herkenbaar zijn.

Verklarend of pragmatisch

Therapeutische experimenten worden in een groot aantal verschillende omstandigheden en met verschillende probleemstellingen uitgevoerd. Dientengevolge zijn er een aantal varianten op het grondthema ontstaan. In verband hiermee kunnen therapeutische experimenten worden onderverdeeld in verklarende en pragmatische trials. In een verklarende trial is men op zoek naar het specifieke effect van het veronderstelde werkzame bestanddeel van de interventie. Daartoe zal deze interventie worden vergeleken met een placebo-interventie die in alles behalve de aanwezigheid van het werkzame bestanddeel aan de interventie gelijk is. Ook is vergelijking mogelijk met een andere interventie, die uitsluitend ten aanzien van het veronderstelde werkzame bestanddeel van de eerstgenoemde interventie verschilt. Verklarende trials in de meest pure vorm worden aangetroffen bij geneesmiddelenonderzoek. De verwante problemen van het vinden van een adequate placebo-interventie en het realiseren van optimale blindering, maken dat verklarende trials op het gebied van de fysiotherapie nog (te) weinig werden gedaan.

De tegenpool wordt gevormd door de pragmatische trial. Hierin worden bepaalde behandelstrategieën met elkaar vergeleken, en niet het geneesmiddel of de therapie afzonderlijk. Het gaat er dan om een vergelijking te maken tussen reële opties in de behandeling van de betreffende patiënten. Het doel van de pragmatische trial is een uitspraak over het effect van een interventie in de praktijk, met daarbij inbegrepen alle aspecten die bij die interventie horen. Omdat niet-behandelen veelal een redelijk alternatief is, zal in een pragmatische trial doorgaans een deel van de patiënten niet worden behandeld. Daarentegen zijn (opzettelijke) placebo-therapiën relatief zeldzaam in de praktijk en derhalve zal in pragmatische trial zelden een placebo worden gebruikt. Daarom zal in een dergelijke trial in de regel geen onderscheid kunnen worden gemaakt tussen het specifieke en het aspecifieke effect (placebo-effect) van een bepaalde behandeling. Maar dat is voor de (para-)medicus practicus dan ook een relatief weinig relevante vraag. Het gaat er immers om of het gewenste effect wordt bereikt, hoe is minder belangrijk.

In het voorbeeld is in feite sprake van een combinatie van een pragmatische en een verklarende trial. De vergelijking tussen manuele therapie en fysiotherapie enerzijds en de placebo-therapie anderzijds, draagt de kenmerken van een verklarende trial. Het beoogde verschil tussen deze interventies zijn de veronderstelde specifieke bestanddelen van manuele therapie, respectievelijk van fysiotherapie. Het pragmatische standpunt wordt echter in

het voorbeeld eveneens ingenomen. Het gaat immers in de onderzoeksopzet om de vergelijking van het effect van drie redelijke verwijsalternatieven voor de huisarts: zelf behandelen, verwijzen naar een fysiotherapeut of naar een manueel therapeut. In feite wordt het verklarende aspect pas vooral interessant, als zou blijken dat deze drie inhoudelijke zo verschillende interventies elkaar qua waargenomen effect niet veel ontlopen. Dan wordt de vraag relevant in hoeverre deze interventies meer effect sorteren dan de placebo-interventie.

Cross-over

Een andere indeling van therapeutische experimenten is die in parallelle en cross-over trials. In het voorgaande is steeds impliciet uitgegaan van een parallelle opzet: geschikte patiënten worden gerandomiseerd en verdeeld over de in de trial te bestuderen interventies. Elke patiënt krijgt daarbij slechts één interventie toegewezen. Dat is niet het geval in een cross-over trial. Daarin krijgen alle patiënten achtereenvolgens alle interventies die met elkaar worden vergeleken. In een cross-over trial bepaalt de randomisatie niet wie welke behandeling krijgt, maar in welke volgorde de interventies bij de betreffende patiënt worden uitgevoerd. Van de cross-over opzet gaat een grote aantrekkingskracht uit. Immers, elke patiënt vormt zo de eigen referentie voor het effect van de verschillende behandelingen. Deze efficiëntie komt ook in de bijbehorende statistische analyse tot uiting. Dientengevolge zijn voor een cross-over experiment relatief weinig patiënten nodig. De cross-over opzet is vooral in farmacologische kringen erg populair (geweest).

Er zijn echter twee voorwaarden waaraan het betreffende therapeutische experiment moet voldoen, wil de cross-over opzet valide zijn.

1. Afwezigheid van een 'carry-over' effect

Wanneer het effect van een interventie van invloed is op de effectiviteit van de erop volgende interventie, is sprake van een carry-over effect. Dit zal met name optreden wanneer de tijdsspanne tussen de interventies ('wash out period') kort is in vergelijking tot de werkingsduur van de interventie, of wanneer de eerdere interventie met enige vertraging werkzaam is.

In het voorbeeld zou van het carry-over effect sprake zijn, wanneer eerst manuele therapie wordt toegepast en het effect hiervan pas later ten tijde van de fysiotherapie zichtbaar wordt. Het probleem is in wezen fundamenteel. Het is ernstig de vraag of de patiënten in de tweede helft van de trial nog wel 'random' over de onderzoeksgroepen zullen zijn verdeeld. Er kunnen tengevolge van de eerste behandeling prognostische verschillen tussen de groepen zijn ontstaan, bijvoorbeeld door selectieve uit-

val of doordat de ervaringen met de eerste behandeling de verwachting ten aanzien van de volgende behandeling selectief beïnvloeden.

2. Het effect van de interventie hangt samen met een reversibele gebeurtenis

Wanneer de effectparameter bijvoorbeeld sterfte, myocardinfarct of conceptie is, kunnen de patiënten waarbij deze gebeurtenis optrad, niet meer aan de volgende interventie worden onderworpen. Wanneer in het voorbeeld met manuele therapie veel patiënten volledig herstellen, ontstaat een vergelijkbare asymmetrie in de onderzoeksopzet. Dergelijke patiënten kunnen immers niet meer aan een volgende interventie worden onderworpen.

Aan de twee voorwaarden voor een valide gebruik van de cross-over opzet zal in veel gevallen niet kunnen worden voldaan. In feite komen alleen interventies in aanmerking die een snel en vooral kortdurend effect bewerkstelligen, waarna de effectparameters weer naar het uitgangsnivo terugkeren. Slaapmiddelen en pijnstillers tegen acute pijn zijn voorbeelden van dergelijke interventies. Het zal duidelijk zijn van dat de cross-over opzet in interventie-onderzoek betreffende de fysiotherapie vrijwel nooit adequaat is.

Uitvoerbaarheid

Interventie-onderzoek in de vorm van een therapeutisch experiment naar aspecten van het fysiotherapeutisch handelen, vereist een grondige voorbereiding en een strakke organisatie. De volgende punten verdienen daarbij bijzondere aandacht.

• Vroegtijdige inschakeling van de (para-) medici die de interventies gaan uitvoeren

Een consensus over de geschikte groep patiënten en een precieze afbakening van de interventies in kwestie staan daarbij centraal. Ook de relevantie van de trial voor het fysiotherapeutisch handelen wordt hiermee gewaarborgd. Tenslotte zal een goed en tijdig overleg met de betrokken (para-)medici motiverend werken, waardoor men zich voor het welslagen van de trial zal inzetten en zich nauwkeurig aan het onderzoeksprotocol zal houden. In het voorbeeld is met de betreffende fysiotherapeuten en manueel therapeuten gedurende een jaar voorafgaand aan het eigenlijke onderzoek ongeveer eens in de zes weken een bijeenkomst belegd.

• Het formuleren van een volledige uitgewerkt onderzoeksprotocol en het ontwerpen van een adequate wijze van de gegevensverzameling

Het is van belang vóór de aanvang van de trial zoveel mogelijk aspecten van het onderzoek in de praktijk uit te proberen. Om begrijpelijke redenen dienen nadat het onderzoek is begonnen, geen veranderingen meer in het protocol te worden aangebracht. In het voorbeeld zullen kleine 'pi-

lot studies' worden gedaan naar de verschillende aspecten van de gegevensverzameling. Bijzondere aandacht wordt daarbij gegeven aan het testen van de meting van de beweeglijkheid van de wervelkolom. Omdat in deze trial, zoals in veel interventie-onderzoek binnen de fysiotherapie, verschillende (para-)medici de interventies uitvoeren, zullen ook de organisatorische kanten van het onderzoek zoveel mogelijk vooraf worden uitgetoetst. Daaronder zijn inbegrepen de standaard formuleren waarmee de gegevens worden verzameld.

• Informed consent procedure

Zorgvuldigheid is geboden bij het recruter van patiënten voor een therapeutisch experiment. Het idee is dat patiënten op basis van vrije wil beslissen over deelname na adequate informatie over het onderzoek te hebben ontvangen. Daarbij is inbegrepen een uitleg van het randomiseren en een vermelding van de kans op placebo-therapie. Het is duidelijk dat er een zekere spanning bestaat tussen de belangen van de individuele patiënt enerzijds, en de wetenschappelijke beantwoording van de vraagstelling anderzijds. Voorwaarden voor het opzetten van de trial zijn, dat geen van de interventies onaanvaardbare risico's met zich meebrengt en dat over de effectiviteit van de vergeleken interventies geen consensus bestaat. Met andere woorden: een therapeutisch experiment richt zich op situaties waarin deskundige (para-)medici het niet eens zijn over de beste behandeling. Veel trials evalueren de waarde van een nieuw geïntroduceerde behandeling, maar vaak blijkt ook dat van gangbare therapieën de werkzaamheid nog onvoldoende is onderzocht. In het voorbeeld is dit het geval. In de informed consent procedure wordt zowel mondeling als schriftelijk uitgelegd dat nog onvoldoende is onderzocht of niet-specifieke rug- en nekklachten beter door de huisarts zelf kunnen worden behandeld, dan wel worden doorverwezen naar een manueel therapeut of een fysiotherapeut. Ook wordt in eenvoudige bewoordingen uitgelegd waarom een kwart van de patiënten, zonder dit zelf te weten, een placebo-behandeling zal krijgen.

Tenslotte

In het bovenstaande is getracht in hoofdlijnen de opzet van het therapeutische experiment te schetsen en deze toe te passen op een voorbeeld uit de fysiotherapie. In het bestek van een artikel kan van een gedetailleerde handleiding geen sprake zijn. Daartoe zijn adequate handboeken beschikbaar (13, 14, 15). Zo is bijvoorbeeld geheel voorbij gegaan aan het fenomeen effect-modificatie. Dit houdt in dat een therapie bijvoorbeeld veel sterker werkzaam is in aanwezigheid van een bepaalde externe factor. Men komt dan in de verleiding, vooral bij afwezigheid van een verschil in werkzaamheid van de bestudeerde

interventies, om een effect in een subgroep van de onderzoekspopulatie te rapporteren. Dit is echter een slechte gewoonte, omdat de kans op een toevalsbevinding aanzienlijk is. Voor een adequate kwantificering van effect-modificatie zijn bijzonder grote aantallen patiënten nodig. Een gedegen uitwerking van de methodologie van de Randomized Controlled Trial alleen, zonder allerlei tips voor de praktische uitvoering, is te vinden in diverse epidemiologische leerboeken (12, 16, 17, 18).

Het therapeutisch experiment is hier gepresenteerd als een paradigma. Het spreekt voor zich dat een dergelijke aanpak soms niet mogelijk of wenselijk is. Zo zijn zeldzame bijwerkingen van bepaalde behandelingen niet of nauwelijks in een therapeutisch experiment onderzoekbaar. Door de zeldzaamheid van de bijwerking zouden immer gigantische aantallen patiënten in de trial moeten worden opgenomen. In 1984 verscheen in het blad *Statistics in Medicine* een aantal bijdragen, die evenwichtig de voor- en nadelen van Randomized Controlled Trials weergeven (19, 20, 21, 22). De centrale boodschap van deze artikelenserie is dat men er goed aan doet, ook daar waar van een therapeutisch experiment wordt afgezien, in interventie-onderzoek steeds de strenge methodologische maatstaven aan te leggen die voor Randomized Controlled Trials worden gehanteerd.

Noot

Het idee voor dit artikel is ontstaan gedurende een cursus 'Principles of intervention research in extramural medicine' die door de Vrije

Universiteit tezamen met de Nederlandse Hartstichting van 3-16 november 1986 werd georganiseerd. De auteurs willen hierbij prof. dr. O.S. Miettinen, prof. dr. J. Lubsen en dr. P. D. Bezemer danken voor het inspirerende onderwijs betreffende de methodologie van Randomized Controlled Trials.

Summary

The importance of effect-evaluation concerning physical therapy is rightly stressed upon from several sides. Such intervention research should have an experimental design. In this article the main points in designing and conducting Randomized Controlled Trials within physical therapy are explained. A recent proposal for such a trial from the authors is taken as an example. Attention is furthermore given to a proper formulation of the research problem, the several components of the effect and the need for randomization and blinding. Furthermore two different classifications of Randomized Controlled Trials are given and more specialized literature on the design of intervention research is mentioned.

Literatuur

1. Apeldoorn, A.T., O.G. Meijer & M. van Oort, *Verwetenschappelijking van de fysiotherapie*. Ned. T. Fysiotherapie 95(1985) 105-111.
2. Vuyk, K., *Verwetenschappelijking en theorievorming in de Fysiotherapie*. Ned. T. Fysiotherapie 95(1985) 46-50.
3. Beeftink, A.D., J. Fennema & J. L. M. van Gestel, *Wetenschappelijk onderzoek in de fysiotherapie*. Ned. T. Fysiotherapie 96(1986) 128-134.
4. Bouter, L.M. & J.M. Keppel Hesselink, *Therapie en effect*. Ned. T. Fysiotherapie 93(1983) 357-360.
5. Bobbert, M., E. Houtkoop, R.H. Rozendal & H.W.R. van Sambeek, *Onderzoek naar effecten van fysiotherapeutisch handelen*. Ned. T. Fysiotherapie 94(1984) 158-160.

6. Aufdemkampe, G., M.A.J. Beijer, O.G. Meijer, H.J.M. Obbens & T.J.A. Terlouw, *Kwaliteit van effectmeting in de fysiotherapie*. Ned. T. Fysiotherapie 95(1985) 123-131.
7. Kuhn, T.S., *The structure of scientific revolutions*. Chicago, 1962.
8. Bouter, L.M. *Het effect van manuele therapie bij lage rugklachten*. Medisch Contact 40 (1985) 357-360.
9. Bouter, L.M. & P.G. Knipschild, *De waarde van manuele therapie en fysiotherapie bij specifieke lage rugpijn*. Tijdschrift voor Sociale Gezondheidszorg 63(1985) 956-957.
10. Miettinen, O.S., R.C. Ellison, G.J. Peckam, E.N. Kraybill, G.L. Breart & A.S. Nadas, *Overall prognosis as the primary criterion of outcome in a clinical trial*. Controlled Clinical Trials 4(1983) 227-237.
11. Miettinen, O.S., *The need for randomization in the study of intended effects*. Statistics in Medicine 2(1983) 267-271.
12. Miettinen, O.S., *Theoretical epidemiology*. New York, 1985.
13. Friedman, L.M., C.D. Furbert & D.L. DeMets, *Fundamentals of clinical trials*. Boston, 1982.
14. Pocock, S.J., *Clinical trials*. Chichester, 1983.
15. Meinert, C.L., *Clinical trials*. New York, 1986.
16. Feinstein, A.R., *Clinical epidemiology*. Philadelphia, 1985.
17. Sackett, D.L., R.B. Haynes & P. Tugwell, *Clinical epidemiology*. Boston, 1985.
18. Ibrahim, M.A., *Epidemiology and health policy*. Rockville, 1985.
19. Brown, B.W., *The randomized clinical trial*. Statistics in Medicine 3(1984) 307-311.
20. Gehan, E.A., *The evaluation of therapies: Historical control studies*. Statistics in Medicine 3(1984) 315-324.
21. Feinstein, A., *The role of observational studies in the evaluation of therapy*. Statistics in Medicine 3(1984) 341-345.
22. Yusuf, S., R. Collins & R. Peto, *Why do we need some large, simple randomized trials*. Statistics in Medicine 3(1984) 409-420.